SEQUENCE LISTING

<110> Lassen, Soren Flensted

<120> Improved proteases and methods for producing them

<130> 10495.204-US

<160> 53

<170> PatentIn version 3.3

<210> 1
<211> 1062
<212> DNA
<213> Nocardiopsis sp. NRRL 18262

<220>
<221> misc_feature
<222> (1)..(495)
<223> Encodes the pro-region shown in positions -165 to -1 of SEQ ID
       NO:43.

<220>
<221> misc_feature
<222> (496)..(1059)
<223> Encodes the mature region shown in positions 1-188 of SEQ ID
       NO:43.

<400> 1
gctactggag cattacctca gtctcctaca cctgaagcag atgcagtatc gatgcaagaa   60

gcattacaac gtgatcttga tcttacatca gctgaagctg aggaattact tgctgcacaa   120

gatacagcct ttgaagttga tgaagctgcc gctgaagcag ctggtgatgc atatggtggt   180

tcagtattcg atactgaatc actcgaactt actgtactag tgaccgatgc agcagctgtt   240

gaagctgttg aagccacagg tgcaggtaca gagctcgtat cttatggtat tgatggatta   300

gatgagatcg tacaagagct taatgcagct gatgccgttc caggtgtagt tggatggtat   360

cctgatgtag caggtgatac tgttgtctta gaagttcttg aaggctctgg agctgatgtt   420

tctggacttt tagcagacgc aggagtcgat gcatccgcgg ttgaagtgac cacgtcagat   480

cagcctgaac tctatgccga tatcattgga ggcctagcgt acacaatggg tggtcgctgc   540

agcgtaggat ttgcagccac aaatgcagct ggacaacctg cttcgtgac agctggacat   600

tgcggccgcg tcggtacaca ggttactatc ggcaatggaa gaggtgtctt tgagcaaagc   660

gtatttcccg ggaatgatgc tgccttcgtt agaggtacgt ccaactttac gcttactaac   720

ttagtatcta gatacaacac tggcggatat gcaactgtag caggtcacaa tcaagcacct   780

```
attggctcta gcgtctgccg ctcagggtcg actacaggat ggcattgtgg aaccattcaa    840

gctagaggtc agagcgtgag ctatcctgaa ggtaccgtaa cgaacatgac tcgtacgact    900

gtatgtgcag aaccaggtga ctctggaggt tcatatatca gcggtacgca agcgcaaggc    960

gttacctcag gtggatccgg taactgtagg acaggtggca caacgttcta ccaggaagtg   1020

acaccgatgg tgaactcttg gggagttaga ctccgtacat aa                       1062
```

<210>   2
<211>   1143
<212>   DNA
<213>   Artificial sequence

<220>
<223>   A synthetic 10R gene (10Rsynt-15) encoding a S2A protease denoted
         "10R" fused by PCR in frame to the signal peptide encoding
         sequence of a heterologous protease, Savinase.

<400>   2
```
atgaagaaac cgttggggaa aattgtcgca agcaccgcac tactcatttc tgttgctttt    60

agttcatcga tcgcatcggc tgctactgga gcattacctc agtctcctac acctgaagca   120

gatgcagtat cgatgcaaga agcattacaa cgtgatcttg atcttacatc agctgaagct   180

gaggaattac ttgctgcaca agatacagcc tttgaagttg atgaagctgc cgctgaagca   240

gctggtgatg catatggtgg ttcagtattc gatactgaat cactcgaact tactgtacta   300

gtgaccgatg cagcagctgt tgaagctgtt gaagccacag gtgcaggtac agagctcgta   360

tcttatggta ttgatggatt agatgagatc gtacaagagc ttaatgcagc tgatgccgtt   420

ccaggtgtag ttggatggta tcctgatgta gcaggtgata ctgttgtctt agaagttctt   480

gaaggctctg gagctgatgt ttctggactt ttagcagacg caggagtcga tgcatccgcg   540

gttgaagtga ccacgtcaga tcagcctgaa ctctatgccg atatcattgg aggcctagcg   600

tacacaatgg gtggtcgctg cagcgtagga tttgcagcca caaatgcagc tggacaacct   660

ggcttcgtga cagctggaca ttgcggccgc gtcggtacac aggttactat cggcaatgga   720

agaggtgtct ttgagcaaag cgtatttccc gggaatgatg ctgccttcgt tagaggtacg   780

tccaacttta cgcttactaa cttagtatct agatacaaca ctggcggata tgcaactgta   840

gcaggtcaca atcaagcacc tattggctct agcgtctgcc gctcagggtc gactacagga   900

tggcattgtg gaaccattca agctagaggt cagagcgtga gctatcctga aggtaccgta   960

acgaacatga ctcgtacgac tgtatgtgca gaaccaggtg actctggagg ttcatatatc   1020

agcggtacgc aagcgcaagg cgttacctca ggtggatccg gtaactgtag gacaggtggc   1080
```

2

acaacgttct accaggaagt gacaccgatg gtgaactctt ggggagttag actccgtaca   1140

taa   1143

<210> 3
<211> 8
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the
       invention.

<400> 3

Gln Ser His Val Gln Ser Ala Pro
1   5

<210> 4
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as
       fusion to protease of the invention.

<400> 4
caatcgcatg ttcaatccgc tcca   24

<210> 5
<211> 4
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the
       invention.

<400> 5

Gln Ser Ala Pro
1

<210> 6
<211> 12
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as
       fusion to protease of the invention.

<400> 6
caatcggctc ct   12

<210> 7
<211> 2
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the
      invention.

<400> 7

Gln Pro
1

<210> 8
<211> 6
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as
      fusion to protease of the invention.

<400> 8
caacca   6

<210> 9
<211> 1
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the
      invention.

<400> 9

Pro
1

<210> 10
<211> 3
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as
      fusion to protease of the invention.

<400> 10
cca   3

<210> 11
<211> 45
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #252639

<400> 11
catgtgcatg tgggtaccgc aacgttcgca gatgctgctg aagag   45

<210> 12
<211> 44
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #251992

<400> 12
catgtgcatg tggtcgaccg attatggagc ggattgaaca tgcg   44

<210> 13
<211> 44
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #179541

<400> 13
gcgttgagac gcgcggccgc gagcgccgtt tggctgaatg atac   44

<210> 14
<211> 43
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #179542

<400> 14
gcgttgagac agctcgagca gggaaaaatg gaaccgcttt ttc   43

<210> 15
<211> 64
<212> DNA
<213> Artificial sequence

```
<220>
<223>  Primer #179539

<400>  15
ccatttgatc agaattcact ggccgtcgtt ttacaaccat tgcggaaaat agtcataggc     60

atcc                                                                  64


<210>  16
<211>  60
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer #179540

<400>  16
ggatccagat ctggtacccg ggtctagagt cgacgcggcg gttcgcgtcc ggacagcaca     60


<210>  17
<211>  37
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer #179154

<400>  17
gttgtaaaac gacggccagt gaattctgat caaatgg                              37


<210>  18
<211>  37
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer #179153

<400>  18
ccgcgtcgac actagacacg ggtacctgat ctagatc                              37


<210>  19
<211>  22
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer #317

<400>  19
tggcgcaatc ggtaccatgg gg                                              22
```

<210> 20
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #139 NotI

<400> 20
catgtgcatg cggccgcatt aacgcgttgc cgcttctgcg   40

<210> 21
<211> 7443
<212> DNA
<213> Artificial sequence

<220>
<223> Sequence of plasmid pMB1508

<400> 21
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca   60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg   120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc   180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc   240

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat   300

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta cgccagggt   360

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cgataaaagt gctttttttg   420

ttgcaattga agaattatta atgttaagct taattaaaga taatatcttt gaattgtaac   480

gcccctcaaa agtaagaact acaaaaaaag aatacgttat atagaaatat gtttgaacct   540

tcttcagatt acaaatatat tcggacggac tctacctcaa atgcttatct aactatagaa   600

tgacatacaa gcacaacctt gaaaatttga aaatataact accaatgaac ttgttcatgt   660

gaattatcgc tgtatttaat tttctcaatt caatatataa tatgccaata cattgttaca   720

agtagaaatt aagacaccct tgatagcctt actataccta acatgatgta gtattaaatg   780

aatatgtaaa tatatttatg ataagaagcg acttatttat aatcattaca tatttttcta   840

ttggaatgat taagattcca atagaatagt gtataaatta tttatcttga aaggagggat   900

gcctaaaaac gaagaacatt aaaaacatat atttgcaccg tctaatggat ttatgaaaaa   960

tcattttatc agtttgaaaa ttatgtatta tggagctctg aaaaaaagga gaggataaag   1020

aatgaagaaa ccgttgggga aaattgtcgc aagcaccgca ctactcattt ctgttgcttt   1080

7

```
tagttcatcg atcgcatcgg ctgctgaaga agcaaaagaa aaatatttaa ttggcttaa     1140

tgagcaggaa gctgtcagtg agtttgtaga acaagtagag gcaaatgacg aggtcgccat     1200

tctctctgag gaagaggaag tcgaaattga attgcttcat gaatttgaaa cgattcctgt     1260

tttatccgtt gagttaagcc cagaagatgt ggacgcgctt gaactcgatc cagcgatttc     1320

ttatattgaa gaggatgcag aagtaacgac aatggcgcaa tcggtaccat ggggtatatc     1380

aacgcgttaa tccgcggata tatagcggcc gcagatctgg gaccaataat aatgactaga     1440

gaagaaagaa tgaagattgt tcatgaaatt aaggaacgaa tattggataa agtgggatat     1500

ttttaaaata tatatttatg ttacagtaat attgactttt aaaaaaggat tgattctaat     1560

gaagaaagca gacaagtaag cctcctaaat tcactttaga taaaaattta ggaggcatat     1620

caaatgaact ttaataaaat tgatttagac aattggaaga gaaaagagat atttaatcat     1680

tatttgaacc aacaaacgac ttttagtata accacagaaa ttgatattag tgtttttatac    1740

cgaaacataa aacaagaagg atataaattt taccctgcat ttattttctt agtgacaagg     1800

gtgataaact caaatacagc ttttagaact ggttacaata gcgacggaga gttaggttat     1860

tgggataagt tagagccact ttatacaatt tttgatggtg tatctaaaac attctctggt     1920

atttggactc ctgtaaagaa tgacttcaaa gagtttatg atttatacct ttctgatgta     1980

gagaaatata atggttcggg gaaattgttt cccaaaacac ctatacctga aaatgctttt     2040

tctctttcta ttattccatg gacttcattt actgggttta acttaaatat caataataat     2100

agtaattacc ttctacccat tattacagca ggaaaattca ttaataaagg taattcaata     2160

tatttaccgc tatctttaca ggtacatcat tctgtttgtg atggttatca tgcaggattg     2220

tttatgaact ctattcagga attgtcagat aggcctaatg actggctttt ataatatgag     2280

ataatgccga ctgtactttt tacagtcggt tttctaacga tacattaata ggtacgaaaa     2340

agcaactttt tttgcgctta aaaccagtca taccaataac ttaagggtaa ctagcctcgc     2400

cggaaagagc gaaaatgcct cacatttgtg ccacctaaaa aggagcgatt tacatatgag     2460

ttatgcagtt tgtagaatgc aaaaagtgaa atcagctgga ctaaaagggg ccgcagagta     2520

gaatggaaaa ggggatcgga aaacaagtat ataggaggag acctatttat ggcttcagaa     2580

aaagacgcag gaaaacagtc agcagtaaag cttgttccat tgcttattac tgtcgctgtg     2640

ggactaatca tctggtttat tcccgctccg tccggacttg aacctaaagc ttggcatttg     2700

tttgcgattt ttgtcgcaac aattatcggc tttatctcca gcccttgcc aatgggtgca      2760

attgcaattt ttgcattggc ggttactgca ctaactggaa cactatcaat tgaggataca     2820
```

8

```
ttaagcggat tcgggaataa gaccatttgg cttatcgtta tcgcattctt tatttcccgg    2880

ggatttatca aaaccggtct cggtgcgaga atttcgtatg tattcgttca gaaattcgga    2940

aaaaaaaccc ttggactttc ttattcactg ctattcagtg atttaatact ttcacctgct    3000

attccaagta atacggcgcg tgcaggaggc attatatttc ctattatcag atcattatcc    3060

gaaacattcg gatcaagccc ggcaaatgga acagagagaa aaatcggtgc attcttatta    3120

aaaaccggtt ttcaggggaa tctgatcaca tctgctatgt tcctgacagc gatggcggcg    3180

aacccgctga ttgccaagct ggcccatgat gtcgcagggg tggacttaac atggacaagc    3240

tgggcaattg ccgcgattgt accgggactt gtaagcttaa tcatcacgcc gcttgtgatt    3300

tacaaactgt atccgccgga aatcaaagaa acaccggatg cggcgaaaat cgcaacagaa    3360

aaactgaaag aaatgggacc gttcaaaaaa tcggagcttt ccatggttat cgtgtttctt    3420

ttggtgcttg tgctgtggat ttttggcggc agcttcaaca tcgacgctac cacaaccgca    3480

ttgatcggtt tggccgttct cttattatca caagttctga cttgggatga tatcaagaaa    3540

gaacagggcg cttgggatac gctcacttgg tttgcggcgc ttgtcatgct cgccaacttc    3600

ttgaatgaat taggcatggt gtcttggttc agtaatgcca tgaaatcatc cgtatcaggg    3660

ttctcttgga ttgtggcatt catcatttta attgttgtgt attattactc tcactatttc    3720

tttgcaagtg cgacagccca catcagtgcg atgtattcag catttttggc tgtcgtcgtg    3780

gcagcgggcg caccgccgct tttagcagcg ctgagcctcg cgttcatcag caacctgttc    3840

gggtcaacga ctcactacgg ttctggagcg gctccggtct tcttcggagc aggctacatc    3900

ccgcaaggca aatggtggtc catcggattt atcctgtcga ttgttcatat catcgtatgg    3960

cttgtgatcg gcggattatg gtggaaagta ctaggaatat ggtagaaaga aaaaggcaga    4020

cgcggtctgc cttttttttat tttcactcct tcgtaagaaa atggattttg aaaaatgaga    4080

aaattccctg tgaaaaatgg tatgatctag gtagaaagga cggctggtgc tgtggtgaaa    4140

aagcggttcc attttttccct gcaaacaaaa ataatggggc tgattgcggc tctgctggtc    4200

tttgtcattg gtgtgctgac cattacgtta gccgttcagc atacacaggg agaacggaga    4260

caggcagagc agctggcggt tcaaacggcg agaaccattt cctatatgcc gccggttaaa    4320

gagctcattg agagaaaaga cggacatgcg gctcagacgc aagaggtcat tgaacaaatg    4380

aaagaacaga ctggtgcgtt tgccatttat gttttgaacg aaaaaggaga cattcgcagc    4440

gcctctggaa aaagcggatt aaagaaactg gagcgcagca gagaaatttt gtttggcggt    4500
```

9

```
tcgcatgttt ctgaaacaaa agcggatgga cgaagagtga tcagagggag cgcgccgatt    4560

ataaaagaac agaagggata cagccaagtg atcggcagcg tgtctgttga ttttctgcaa    4620

acggagacag agcaaagcat caaaaagcat ttgagaaatt tgagtgtgat tgctgtgctt    4680

gtactgctgc tcggatttat tggcgccgcc gtgctggcga aaagcatcag aaaggatacg    4740

ctcgggcttg aaccgcatga gatcgcggct ctatatcgtg agaggaacgc aatgcttttc    4800

gcgattcgag aagggattat tgccaccaat cgtgaaggcg tcgtcaccat gatgaacgta    4860

tcggcggccg agatgctgaa gctgcccgag cctgtgatcc atcttcctat agatgacgtc    4920

atgccgggag cagggctgat gtctgtgctt gaaaaaggag aaatgctgcc gaaccaggaa    4980

gtaagcgtca acgatcaagt gtttattatc aatacgaaag tgatgaatca aggcgggcag    5040

gcgtatggga ttgtcgtcag cttcagggag aaaacagagc tgaagaagct gatcgacaca    5100

ttgacagagg ttcgcaaata ttcagaggat ctcagggcgc agactcatga attttcaaat    5160

aagctttatg cgattttagg gctgcgtcga cctgcaggca tgcaagcttg gcgtaatcat    5220

ggtcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag    5280

ccggaagcat aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg    5340

cgttgcgctc actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa    5400

tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca    5460

ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg    5520

taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc    5580

agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc    5640

cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac    5700

tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc    5760

tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata    5820

gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc    5880

acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca    5940

acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag    6000

cgaggtatgt aggcggtgct acagagttct tgaagtggtg cctaactac ggctacacta    6060

gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg    6120

gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc    6180

agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt    6240
```

10

```
ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa    6300

ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat    6360

atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga    6420

tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac    6480

gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg    6540

caccggattt atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg    6600

caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt    6660

cgccagttaa tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct    6720

cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat    6780

cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta    6840

agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca    6900

tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat    6960

agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac    7020

atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa    7080

ggatcttacc gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt    7140

cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg    7200

caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc cttttttcaat    7260

attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt    7320

agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacgtct    7380

aagaaaccat tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc    7440

gtc                                                                  7443


<210>    22
<211>    5718
<212>    DNA
<213>    Artificial sequence

<220>
<223>    Sequence of MB1510 genomic integration region

<400>    22
gagcgccgtt tggctgaatg atacaacagt ctcacttcct tactgcgtct ggttgcaaaa    60

acgaagaagc aaggattccc ctcgcttctc atttgtccta tttattatac actttttttaa   120
```

11

```
gcacatcttt ggcgcttgtt tcactagact tgatgcctct gaatcttgtc caagtgtcac    180

ggtccgcatc atagacttgt ccattttca ccgctttgag attttccag agcgggttcg    240

ttttccactc atctacaatg gttttgcctt cgttggctga gatgaacaaa atatcaggat    300

cgattttgct caattgctca aggctgacct cttgataggc gttatctgac ttcacagcgt    360

gtgtaaagcc tagcatttta aagatttctc cgtcatagga tgatgatgta tgaagctgga    420

aggaatccgc tcttgcaacg ccgagaacga tgttgcggtt ttcatctttc ggaagttcgg    480

cttttagatc gttgatgact tttttgtgct cggcaagctt ttcttttcct tcatcttctt    540

tatttaatgc tttagcaatg gtcgtaaagc tgtcgatcgt ttcgtcatat gtcgcttcac    600

ggctttttaa ttcaatcgtc ggggcgattt ttttcagctg tttataaatg tttttatggc    660

gctcagcgtc agcgatgatt aaatcaggct tcaaggaact gatgacctca agattgggtt    720

cgctgcgtgt gcctacagat gtgtaatcaa tggagctgcc gacaagcttt ttaatcatat    780

cttttttgtt gtcatctgcg atgcccaccg gcgtaatgcc gagattgtga acggcatcca    840

agaatgaaag ctcaagcaca accacccgct taggtgtgcc gcttactgtc gtttttcctt    900

cttcgtcatg gatcactctg gaatccttag actcgctttt gccgcttccg ttgttattct    960

ggcttgatga acagccggat acaatgaggc aggcgagcaa taaaacactc atgatggcaa   1020

tcaacttgtt agaataggtg cgcatgtcat tcttcctttt ttcagattta gtaatgagaa   1080

tcattatcac atgtaacact ataatagcat ggcttatcat gtcaatattt ttttagtaaa   1140

gaaagctgcg tttttactgc tttctcatga aagcatcatc agacacaaat aagtggtatg   1200

cagcgttacc gtgtcttcga dacaaaaacg catgggcgtt ggctttagag gtttcgaaca   1260

tatcagcagt gacataagga aggagagtgc tgagataacc ggacaatttc ttttctattt   1320

catctgttag tgcaaattca atgtcgccga tattcatgat aatcgagaaa acaaagtcga   1380

tatcgatatg aaaatgttcc tcggcaaaaa ccgcaagctc gtgaattcct ggtgaacatc   1440

cggcacgctt atggaaaatc tgtttgacta aatcactcac aatccaagca ttgtattgct   1500

gttctggtga aaagtattgc attagacata cctcctgctc gtacggataa aggcagcgtt   1560

tcatggtcgt gtgctccgtg cagcggcttc tccttaattt tgattttttct gaaaataggt   1620

cccgttccta tcactttacc atggacggaa aacaaatagc tactaccatt cctcctgttt   1680

ttctcttcaa tgttctggaa tctgtttcag gtacagacga tcgggtatga aagaaatata   1740

gaaaacatga aggaggaata tcgacatgaa accagttgta aaagagtata caaatgacga   1800

acagctcatg aaagatgtag aggaattgca gaaaatgggt gttgcgaaag aggatgtata   1860
```

12

```
cgtcttagct cacgacgatg acagaacgga acgcctggct gacaacacga acgccaacac   1920

gatcggagcc aaagaaacag gtttcaagca cgcggtggga aatatcttca ataaaaaagg   1980

agacgagctc cgcaataaaa ttcacgaaat cggttttttct gaagatgaag ccgctcaatt   2040

tgaaaaacgc ttagatgaag gaaaagtgct tctctttgtg acagataacg aaaaagtgaa   2100

agcttgggca taaagcaagg aaaaaaccaa aaggccaatg tcggccttttt ggtttttttg   2160

cggtctttgc ggtgggattt tgcagaatgc cgcaatagga tagcggaaca ttttcggttc   2220

tgaatgtccc tcaatttgct attatatttt tgtgataaat tggaataaaa tctcacaaaa   2280

tagaaaatgg gggtacatag tggatgaaaa aagtgatgtt agctacggct ttgtttttag   2340

gattgactcc agctggcgcg aacgcagctg atttaggcca ccagacgttg ggatccaatg   2400

atggctgggg cgcgtactcg accggcacga caggcggatc aaaagcatcc tcctcaaatg   2460

tgtataccgt cagcaacaga aaccagcttg tctcggcatt agggaaggaa acgaacacaa   2520

cgccaaaaat catttatatc aagggaacga ttgacatgaa cgtggatgac aatctgaagc   2580

cgcttggcct aaatgactat aaagatccgg agtatgattt ggacaaatat ttgaaagcct   2640

atgatcctag cacatggggc aaaaaagagc cgtcgggaac acaagaagaa gcgagagcac   2700

gctctcagaa aaaccaaaaa gcacgggtca tggtggatat ccctgcaaac acgacgatcg   2760

tcggttcagg gactaacgct aaagtcgtgg gaggaaactt ccaaatcaag agtgataacg   2820

tcattattcg caacattgaa ttccaggatg cctatgacta ttttccgcaa tggttgtaaa   2880

acgacggcca gtgaattctg atcaaatggt tcagtgagag cgaagcgaac acttgatttt   2940

ttaattttct atctttttata ggtcattaga gtatacttat ttgtcctata aactatttag   3000

cagcataata gatttattga ataggtcatt taagttgagc atattagagg aggaaaatct   3060

tggagaaata tttgaagaac ccgagatcta gatcaggtac cgcaacgttc gcagatgctg   3120

ctgaagagat tattaaaaag ctgaaagcaa aaggctatca attggtaact gtatctcagc   3180

ttgaagaagt gaagaagcag agaggctatt gaataaatga gtagaaagcg ccatatcggc   3240

gcttttcttt tggaagaaaa tatagggaaa atggtacttg ttaaaaattc ggaatatttta   3300

tacaatatca tatgtatcac attgaaagga ggggcctgct gtccagactg tccgctgtgt   3360

aaaaataagg aataaagggg ggttgacatt attttactga tatgtataat ataatttgta   3420

taagaaaatg gagggggccct cgaaacgtaa gatgaaacct tagataaaag tgcttttttt   3480

gttgcaattg aagaattatt aatgttaagc ttaattaaag ataatatctt tgaattgtaa   3540
```

13

```
cgcccctcaa aagtaagaac tacaaaaaaa gaatacgtta tatagaaata tgtttgaacc   3600

ttcttcagat tacaaatata ttcggacgga ctctacctca aatgcttatc taactataga   3660

atgacataca agcacaacct tgaaaatttg aaaatataac taccaatgaa cttgttcatg   3720

tgaattatcg ctgtatttaa ttttctcaat tcaatatata atatgccaat acattgttac   3780

aagtagaaat taagacaccc ttgatagcct tactatacct aacatgatgt agtattaaat   3840

gaatatgtaa atatatttat gataagaagc gacttattta taatcattac atattttct    3900

attggaatga ttaagattcc aatagaatag tgtataaatt atttatcttg aaaggaggga   3960

tgcctaaaaa cgaagaacat taaaaacata tatttgcacc gtctaatgga tttatgaaaa   4020

atcattttat cagtttgaaa attatgtatt atggagctct gaaaaaaagg agaggataaa   4080

gagaaaaggg gatcggaaaa caagtatata ggaggagacc tatttatggc ttcagaaaaa   4140

gacgcaggaa aacagtcagc agtaaagctt gttccattgc ttattactgt cgctgtggga   4200

ctaatcatct ggtttattcc cgctccgtcc ggacttgaac ctaaagcttg gcatttgttt   4260

gcgatttttg tcgcaacaat tatcggcttt atctccaagc ccttgccaat gggtgcaatt   4320

gcaatttttg cattggcggt tactgcacta actggaacac tatcaattga ggatacatta   4380

agcggattcg ggaataagac catttggctt atcgttatcg cattctttat ttcccgggga   4440

tttatcaaaa ccggtctcgg tgcgagaatt tcgtatgtat tcgttcagaa attcggaaaa   4500

aaaacccttg gactttctta ttcactgcta ttcagtgatt taatactttc acctgctatt   4560

ccaagtaata cggcgcgtgc aggaggcatt atatttccta ttatcagatc attatccgaa   4620

acattcggat caagcccggc aaatggaaca gagagaaaaa tcggtgcatt cttattaaaa   4680

accggttttc aggggaatct gatcacatct gctatgttcc tgacagcgat ggcggcgaac   4740

ccgctgattg ccaagctggc ccatgatgtc gcaggggtgg acttaacatg gacaagctgg   4800

gcaattgccg cgattgtacc gggacttgta agcttaatca tcacgccgct tgtgatttac   4860

aaactgtatc cgccggaaat caaagaaaca ccggatgcgg cgaaaatcgc aacagaaaaa   4920

ctgaaagaaa tgggaccgtt caaaaaatcg gagctttcca tggttatcgt gtttcttttg   4980

gtgcttgtgc tgtggatttt tggcggcagc ttcaacatcg acgctaccac aaccgcattg   5040

atcggtttgg ccgttctctt attatcacaa gttctgactt gggatgatat caagaaagaa   5100

cagggcgctt gggatacgct cacttggttt gcggcgcttg tcatgctcgc caacttcttg   5160

aatgaattag gcatggtgtc ttggttcagt aatgccatga aatcatccgt atcagggttc   5220

tcttggattg tggcattcat cattttaatt gttgtgtatt attactctca ctatttcttt   5280
```

14

```
gcaagtgcga cagcccacat cagtgcgatg tattcagcat ttttggctgt cgtcgtggca     5340

gcgggcgcac cgccgctttt agcagcgctg agcctcgcgt tcatcagcaa cctgttcggg     5400

tcaacgactc actacggttc tggagcggct ccggtcttct tcggagcagg ctacatcccg     5460

caaggcaaat ggtggtccat cggatttatc ctgtcgattg ttcatatcat cgtatggctt     5520

gtgatcggcg gattatggtg gaaagtacta ggaatatggt agaaagaaaa aggcagacgc     5580

ggtctgcctt ttttattttt cactccttcg taagaaaatg gattttgaaa aatgagaaaa     5640

ttccctgtga aaaatggtat gatctaggta gaaaggacgg ctggtgctgt ggtgaaaaag     5700

cggttccatt tttccctg                                                   5718
```

```
<210>   23
<211>   27
<212>   DNA
<213>   Artificial sequence

<220>
<223>   Primer 1605

<400>   23
gacggccagt gaattcgata aaagtgc                                              27
```

```
<210>   24
<211>   42
<212>   DNA
<213>   Artificial sequence

<220>
<223>   Primer 1606


<220>
<221>   misc_feature
<222>   (13)..(13)
<223>   n is a, c, g, or t

<220>
<221>   misc_feature
<222>   (16)..(16)
<223>   n is a, c, g, or t

<400>   24
ccagatctct atnktnktgt acggagtcta actccccaag ag                             42
```

```
<210>   25
<211>   1112
<212>   DNA
<213>   Nocardiopsis dassonvillei DSM 43235
```

```
<400>  25
gctttttagtt catcgatcgc atcggctgct ccggcccccg tcccccagac ccccgtcgcc      60

gacgacagcg ccgccagcat gaccgaggcg ctcaagcgcg acctcgacct cacctcggcc     120

gaggccgagg agcttctctc ggcgcaggaa gccgccatcg agaccgacgc cgaggccacc     180

gaggccgcgg gcgaggccta cggcggctca ctgttcgaca ccgagaccct cgaactcacc     240

gtgctggtca ccgacgcctc cgccgtcgag gcggtcgagg ccaccggagc ccaggccacc     300

gtcgtctccc acggcaccga gggcctgacc gaggtcgtgg aggacctcaa cggcgccgag     360

gttcccgaga gcgtcctcgg ctggtacccg gacgtggaga gcgacaccgt cgtggtcgag     420

gtgctggagg gctccgacgc cgacgtcgcc gccctgctcg ccgacgccgg tgtggactcc     480

tcctcggtcc gggtggagga ggccgaggag gccccgcagg tctacgccga catcatcggc     540

ggcctggcct actacatggg cggccgctgc tccgtcggct cgccgcgac caacagcgcc     600

ggtcagcccg gtttcgtcac cgccggccac tgcggcaccg tcggcaccgg cgtgaccatc     660

ggcaacggca ccggcacctt ccagaactcg gtcttccccg gcaacgacgc cgccttcgtc     720

cgcggcacct ccaacttcac cctgaccaac ctggtctcgc gctacaactc cggcggctac     780

cagtcggtga ccggtaccag ccaggccccg gccggctcgg ccgtgtgccg ctccggctcc     840

accaccggct ggcactgcgg caccatccag gcccgcaacc agaccgtgcg ctacccgcag     900

ggcaccgtct actcgctcac ccgcaccaac gtgtgcgccg agcccggcga ctccggcggt     960

tcgttcatct ccggctcgca ggcccagggc gtcacctccg gcggctccgg caactgctcc    1020

gtcggcggca cgacctacta ccaggaggtc accccgatga tcaactcctg gggtgtcagg    1080

atccggacct aatcgcatgt tcaatccgct cc                                  1112


<210>  26
<211>  48
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer 1423

<400>  26
gctttttagtt catcgatcgc atcggctgct ccggcccccg tcccccag                  48


<210>  27
<211>  45
<212>  DNA
<213>  Artificial sequence
```

16

<220>
<223> Primer 1475

<400> 27
ggagcggatt gaacatgcga ttaggtccgg atcctgacac cccag   45

<210> 28
<211> 354
<212> PRT
<213> Nocardiopsis dassonvillei DSM 43235

<220>
<221> PROPEP
<222> (1)..(166)

<220>
<221> mat_peptide
<222> (167)..(354)

<400> 28

```
Ala Pro  Ala Pro Val Pro Gln  Thr Pro Val Ala Asp  Asp Ser Ala
    -165                -160                -155

Ala Ser  Met Thr Glu Ala Leu  Lys Arg Asp Leu Asp  Leu Thr Ser
    -150                -145                -140

Ala Glu  Ala Glu Glu Leu Leu  Ser Ala Gln Glu Ala  Ala Ile Glu
    -135                -130                -125

Thr Asp  Ala Glu Ala Thr Glu  Ala Ala Gly Glu Ala  Tyr Gly Gly
    -120                -115                -110

Ser Leu  Phe Asp Thr Glu Thr  Leu Glu Leu Thr Val  Leu Val Thr Asp
    -105                -100                -95

Ala Ser Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val
-90               -85                -80                -75

Val Ser His Gly Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn
              -70                -65                -60

Gly Ala Glu Val Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu
              -55                -50                -45

Ser Asp Thr Val Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val
```

```
          -40                      -35                       -30

Ala Ala Leu Leu Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val
    -25                 -20                 -15

Glu Glu Ala Glu Glu Ala Pro Gln Val Tyr Ala Asp Ile Ile Gly Gly
-10                  -5                  -1  1                   5

Leu Ala Tyr Tyr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr
            10                  15                  20

Asn Ser Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr
            25                  30                  35

Val Gly Thr Gly Val Thr Ile Gly Asn Gly Thr Gly Thr Phe Gln Asn
    40                  45                  50

Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn
55                  60                  65                  70

Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Gln
            75                  80                  85

Ser Val Thr Gly Thr Ser Gln Ala Pro Ala Gly Ser Ala Val Cys Arg
            90                  95                  100

Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn
            105                 110                 115

Gln Thr Val Arg Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr
    120                 125                 130

Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly
135                 140                 145                 150

Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Val
            155                 160                 165

Gly Gly Thr Thr Tyr Tyr Gln Glu Val Thr Pro Met Ile Asn Ser Trp
            170                 175                 180

Gly Val Arg Ile Arg Thr
            185
```

18

<210> 29
<211> 498
<212> DNA
<213> Nocardiopsis dassonvillei DSM 43235

<400> 29
gctccggccc ccgtccccca gaccccgtc gccgacgaca gcgccgccag catgaccgag   60

gcgctcaagc gcgacctcga cctcacctcg gccgaggccg aggagcttct ctcggcgcag   120

gaagccgcca tcgagaccga cgccgaggcc accgaggccg cgggcgaggc ctacggcggc   180

tcactgttcg acaccgagac cctcgaactc accgtgctgg tcaccgacgc ctccgccgtc   240

gaggcggtcg aggccaccgg agcccaggcc accgtcgtct cccacggcac cgagggcctg   300

accgaggtcg tggaggacct caacggcgcc gaggttcccg agagcgtcct cggctggtac   360

ccggacgtgg agagcgacac cgtcgtggtc gaggtgctgg agggctccga cgccgacgtc   420

gccgccctgc tcgccgacgc cggtgtggac tcctcctcgg tccgggtgga ggaggccgag   480

gaggccccgc aggtctac   498

<210> 30
<211> 166
<212> PRT
<213> Nocardiopsis dassonvillei DSM 43235

<400> 30

Ala Pro Ala Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala Ala
1   5   10   15

Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asp Leu Thr Ser Ala Glu
                20   25   30

Ala Glu Glu Leu Leu Ser Ala Gln Glu Ala Ala Ile Glu Thr Asp Ala
            35   40   45

Glu Ala Thr Glu Ala Ala Gly Glu Ala Tyr Gly Gly Ser Leu Phe Asp
        50   55   60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65   70   75   80

Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val Val Ser His Gly
                85   90   95

19

```
Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn Gly Ala Glu Val
            100                 105                 110

Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu Ser Asp Thr Val
            115                 120                 125

Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val Ala Ala Leu Leu
    130                 135                 140

Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val Glu Glu Ala Glu
145                 150                 155                 160

Glu Ala Pro Gln Val Tyr
                165


<210>   31
<211>   1146
<212>   DNA
<213>   Artificial sequence

<220>
<223>   The DNA sequence coding for the pro-region of SEQ ID NO: 29 fused
        in frame to A1918L2 protease tail-variant encoding gene; whole
        construct: 10R(proA1918L2).

<400>   31
atgaagaaac cgttggggaa aattgtcgca agcaccgcac tactcatttc tgttgctttt    60

agttcatcga tcgcatcggc tgctccggcc cccgtccccc agacccccgt cgccgacgac   120

agcgccgcca gcatgaccga ggcgctcaag cgcgacctcg acctcacctc ggccgaggcc   180

gaggagcttc tctcggcgca ggaagccgcc atcgagaccg acgccgaggc caccgaggcc   240

gcgggcgagg cctacggcgg ctcactgttc gacaccgaga ccctcgaact caccgtgctg   300

gtcaccgacg cctccgccgt cgaggcggtc gaggccaccg agcccaggc caccgtcgtc    360

tcccacggca ccgagggcct gaccgaggtc gtggaggacc tcaacggcgc cgaggttccc   420

gagagcgtcc tcggctggta cccggacgtg gagagcgaca ccgtcgtggt cgaggtgctg   480

gagggctccg acgccgacgt cgccgccctg ctcgccgacg ccggtgtgga ctcctcctcg   540

gtccgggtgg aggaggccga ggaggccccg caggtctatg ccgatatcat tggaggccta   600

gcgtacacaa tgggtggtcg ctgcagcgta ggatttgcag ccacaaatgc agctggacaa   660

cctggcttcg tgacagctgg acattgcggc cgcgtcggta cacaggttac tatcggcaat   720

ggaagaggtg tctttgagca aagcgtattt cccgggaatg atgctgcctt cgttagaggt   780
```

```
acgtccaact ttacgcttac taacttagta tctagataca acactggcgg atatgcaact    840

gtagcaggtc acaatcaagc acctattggc tctagcgtct gccgctcagg gtcgactaca    900

ggatggcatt gtggaaccat tcaagctaga ggtcagagcg tgagctatcc tgaaggtacc    960

gtaacgaaca tgactcgtac gactgtatgt gcagaaccag gtgactctgg aggttcatat   1020

atcagcggta cgcaagcgca aggcgttacc tcaggtggat ccggtaactg taggacaggt   1080

ggcacaacgt tctaccagga agtgacaccg atggtgaact cttggggagt tagactccgt   1140

acataa                                                              1146
```

<210> 32
<211> 1068
<212> DNA
<213> Nocardiopsis Alba DSM 15647

<400> 32
```
gcgaccggcc ccctcccccca gtcccccacc ccggatgaag ccgaggccac caccatggtc     60

gaggccctcc agcgcgacct cggcctgtcc ccctctcagg ccgacgagct cctcgaggcg    120

caggccgagt ccttcgagat cgacgaggcc gccaccgcgg ccgcagccga ctcctacggc    180

ggctccatct tcgacaccga cagcctcacc ctgaccgtcc tggtcaccga cgcctccgcc    240

gtcgaggcgg tcgaggccgc cggcgccgag gccaaggtgg tctcgcacgg catggagggc    300

ctggaggaga tcgtcgccga cctgaacgcg gccgacgctc agcccggcgt cgtgggctgg    360

taccccgaca tccactccga cacggtcgtc ctcgaggtcc tcgagggctc cggtgccgac    420

gtggactccc tgctcgccga cgccggtgtg gacaccgccg acgtcaaggt ggagagcacc    480

accgagcagc ccgagctgta cgccgacatc atcggcggtc tcgcctacac catgggtggg    540

cgctgctcgg tcggcttcgc ggccaccaac gcctccggcc agcccgggtt cgtcaccgcc    600

ggccactgcg gcaccgtcgg caccccggtc agcatcggca acggccaggg cgtcttcgag    660

cgttccgtct ccccggcaa cgactccgcc ttcgtccgcg gcacctcgaa cttcaccctg    720

accaacctgg tcagccgcta caacaccggt ggttacgcga ccgtctccgg ctcctcgcag    780

gcggcgatcg gctcgcagat ctgccgttcc ggctccacca ccggctggca ctgcggcacc    840

gtccaggccc gcggccagac ggtgagctac ccccagggca ccgtgcagaa cctgacccgc    900

accaacgtct gcgccgagcc cggtgactcc ggcggctcct tcatctccgg cagccaggcc    960

cagggcgtca cctccggtgg ctccggcaac tgctccttcg gtggcaccac ctactaccag   1020

gaggtcaacc cgatgctgag cagctggggt ctgaccctgc gcacctga              1068
```

21

<210> 33
<211> 355
<212> PRT
<213> Nocardiopsis Alba DSM 15647

<220>
<221> PROPEP
<222> (1)..(167)

<220>
<221> mat_peptide
<222> (168)..(355)

<400> 33

```
Ala Thr Gly  Pro Leu Pro Gln Ser  Pro Thr Pro Asp Glu  Ala Glu
        -165                 -160                 -155


Ala Thr Thr  Met Val Glu Ala Leu  Gln Arg Asp Leu Gly  Leu Ser
        -150                 -145                 -140


Pro Ser Gln  Ala Asp Glu Leu Leu  Glu Ala Gln Ala Glu  Ser Phe
        -135                 -130                 -125


Glu Ile Asp  Glu Ala Ala Thr Ala  Ala Ala Ala Asp Ser  Tyr Gly
        -120                 -115                 -110


Gly Ser Ile  Phe Asp Thr Asp Ser  Leu Thr Leu Thr Val  Leu Val. Thr
        -105                 -100                 -95


Asp Ala Ser Ala Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys
     -90             -85             -80


Val Val Ser His Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu
-75             -70             -65             -60


Asn Ala Ala Asp Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile
             -55             -50             -45


His Ser Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp
             -40             -35             -30


Val Asp Ser Leu Leu Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys
        -25             -20             -15
```

22

```
Val Glu Ser Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly
    -10                 -5              -1   1                 5

Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala
              10                  15              20

Thr Asn Ala Ser Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly
              25                  30              35

Thr Val Gly Thr Pro Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu
              40                  45              50

Arg Ser Val Phe Pro Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser
    55                  60                  65

Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr
70                  75                  80                  85

Ala Thr Val Ser Gly Ser Ser Gln Ala Ala Ile Gly Ser Gln Ile Cys
              90                  95                  100

Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Val Gln Ala Arg
              105                 110                 115

Gly Gln Thr Val Ser Tyr Pro Gln Gly Thr Val Gln Asn Leu Thr Arg
              120                 125                 130

Thr Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser
    135                 140                 145

Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser
150                 155                 160                 165

Phe Gly Gly Thr Thr Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser
              170                 175                 180

Trp Gly Leu Thr Leu Arg Thr
              185
```

<220>
<223> Primer 1421

<400> 34
gttcatcgat cgcatcggct gcgaccggcc ccctccccca gtc   43

<210> 35
<211> 31
<212> DNA
<213> Artificial sequence

<220>
<223> Primer 1604

<400> 35
gcggatccta tcaggtgcgc agggtcagac c   31

<210> 36
<211> 1062
<212> DNA
<213> Nocardiopsis prasina DSM 15648

<400> 36
gccaccggac cgctcccccca gtcacccacc ccggaggccg acgccgtctc catgcaggag   60

gcgctccagc gcgacctcgg cctgaccccg cttgaggccg atgaactgct ggccgcccag   120

gacaccgcct tcgaggtcga cgaggccgcg gccgcggccg ccggggacgc ctacggcggc   180

tccgtcttcg acaccgagac cctggaactg accgtcctgg tcaccgacgc cgcctcggtc   240

gaggctgtgg aggccaccgg cgcgggtacc gaactcgtct cctacggcat cgagggcctc   300

gacgagatca tccaggatct caacgccgcc gacgccgtcc ccggcgtggt cggctggtac   360

ccggacgtgg cgggtgacac cgtcgtcctg gaggtcctgg agggttccgg agccgacgtg   420

agcggcctgc tcgccgacgc cggcgtggac gcctcggccg tcgaggtgac cagcagtgcg   480

cagcccgagc tctacgccga catcatcggc ggtctggcct acaccatggg cggccgctgt   540

tcggtcggat cgcggccac caacgccgcc ggtcagcccg gattcgtcac cgccggtcac   600

tgtggccgcg tgggcaccca ggtgagcatc ggcaacggcc agggcgtctt cgagcagtcc   660

atcttcccgg gcaacgacgc cgccttcgtc cgcggcacgt ccaacttcac gctgaccaac   720

ctggtcagcc gctacaacac cggcggttac gccaccgtcg ccggccacaa ccaggcgccc   780

atcggctcct ccgtctgccg ctccggctcc accaccggct ggcactgcgg caccatccag   840

gcccgcggcc agtcggtgag ctaccccgag ggcaccgtca ccaacatgac ccggaccacc   900

gtgtgcgccg agcccggcga ctccggcggc tcctacatct ccggcaacca ggcccagggc   960

gtcacctccg gcggctccgg caactgccgc accggcggga ccaccttcta ccaggaggtc   1020

accccccatgg tgaactcctg gggcgtccgt ctccggacct aa   1062

<210> 37
<211> 353
<212> PRT
<213> Nocardiopsis prasina DSM 15648

<220>
<221> PROPEP
<222> (1)..(165)

<220>
<221> mat_peptide
<222> (166)..(353)

<400> 37

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala
-165   -160   -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro
-150   -145   -140

Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu
-135   -130   -125

Val Asp Glu Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly
-120   -115   -110

Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp
-105   -100   -95   -90

Ala Ala Ser Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu
              -85   -80   -75

Val Ser Tyr Gly Ile Glu Gly Leu Asp Glu Ile Ile Gln Asp Leu Asn
              -70   -65   -60

Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala
         -55   -50   -45

Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val
      -40   -35   -30

25

```
Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val
-25             -20             -15             -10

Thr Ser Ser Ala Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu
            -5              -1  1               5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn
        10              15              20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val
    25              30              35

Gly Thr Gln Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu Gln Ser
40              45              50              55

Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe
            60              65              70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr
        75              80              85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser
    90              95              100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln
    105             110             115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr
120             125             130             135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn
            140             145             150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly
        155             160             165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
        170             175             180

Val Arg Leu Arg Thr
    185
```

26

```
<210>  38
<211>  43
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer 1346

<400>  38
gttcatcgat cgcatcggct gccaccggac cgctccccca gtc                        43


<210>  39
<211>  38
<212>  DNA
<213>  Artificial sequence

<220>
<223>  Primer 1602

<400>  39
gcggatccta ttaggtccgg agacggacgc cccaggag                             38


<210>  40
<211>  1062
<212>  DNA
<213>  Nocardiopsis prasina DSM 15649

<400>  40
gccaccggac cactccccca gtcacccacc ccggaggccg acgccgtctc catgcaggag      60

gcgctccagc gcgacctcgg cctgaccccg cttgaggccg atgaactgct ggccgcccag     120

gacaccgcct tcgaggtcga cgaggccgcg ccgaggccg ccggtgacgc ctacggcggc      180

tccgtcttcg acaccgagac cctggaactg accgtcctgg tcaccgactc cgccgcggtc     240

gaggcggtgg aggccaccgg cgccgggacc gaactggtct cctacggcat cacgggcctc     300

gacgagatcg tcgaggagct caacgccgcc gacgccgttc ccggcgtggt cggctggtac     360

ccggacgtcg cgggtgacac cgtcgtgctg gaggtcctgg agggttccgg cgccgacgtg     420

ggcggcctgc tcgccgacgc cggcgtggac gcctcggcgg tcgaggtgac caccaccgag     480

cagcccgagc tgtacgccga catcatcggc ggtctggcct acaccatggg cggccgctgt     540

tcggtcggct cgcggccac caacgccgcc ggtcagcccg ggttcgtcac cgccggtcac      600

tgtggccgcg tgggcaccca ggtgaccatc ggcaacggcc ggggcgtctt cgagcagtcc     660

atcttcccgg gcaacgacgc cgccttcgtc cgcggaacgt ccaacttcac gctgaccaac     720

ctggtcagcc gctacaacac cggcggctac gccaccgtcg ccggtcacaa ccaggcgccc     780

atcggctcct ccgtctgccg ctccggctcc accaccggtt ggcactgcgg caccatccag     840
```

27

gcccgcggcc agtcggtgag ctaccccgag ggcaccgtca ccaacatgac gcggaccacc   900

gtgtgcgccg agcccggcga ctccggcggc tcctacatct ccggcaacca gcccagggc   960

gtcacctccg gcggctccgg caactgccgc accggcggga ccaccttcta ccaggaggtc   1020

acccccatgg tgaactcctg gggcgtccgt ctccggacct aa   1062

<210> 41
<211> 353
<212> PRT
<213> Nocardiopsis prasina DSM 15649

<220>
<221> PROPEP
<222> (1)..(165)

<220>
<221> mat_peptide
<222> (166)..(353)

<400> 41

Ala   Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala
-165   -160   -155

Val   Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro
-150   -145   -140

Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu
-135   -130   -125

Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly
-120   -115   -110

Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp
-105   -100   -95   -90

Ser Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu
               -85   -80   -75

Val Ser Tyr Gly Ile Thr Gly Leu Asp Glu Ile Val Glu Glu Leu Asn
          -70   -65   -60

Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala
          -55   -50   -45

```
Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val
    -40             -35             -30

Gly Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val
    -25             -20             -15             -10

Thr Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu
                -5              -1  1               5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn
        10              15              20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val
    25              30              35

Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser
40              45              50              55

Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe
            60              65              70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr
        75              80              85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser
    90              95              100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln
    105             110             115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr
120             125             130             135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn
            140             145             150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly
            155             160             165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
        170             175             180
```

29

Val Arg Leu Arg Thr
    185

<210> 42
<211> 43
<212> DNA
<213> Artificial sequence

<220>
<223> Primer 1603

<400> 42
gttcatcgat cgcatcggct gccaccggac cactcccca gtc   43

<210> 43
<211> 353
<212> PRT
<213> Nocardiopsis sp. NRRL 18262

<220>
<221> PROPEP
<222> (1)..(165)

<220>
<221> mat_peptide
<222> (166)..(1059)

<400> 43

Ala   Thr Gly Ala Leu Pro   Gln Ser Pro Thr Pro   Glu Ala Asp Ala
-165   -160.   -155

Val   Ser Met Gln Glu Ala   Leu Gln Arg Asp Leu   Asp Leu Thr Ser
-150   -145   -140

Ala   Glu Ala Glu Glu Leu   Leu Ala Ala Gln Asp   Thr Ala Phe Glu
-135   -130   -125

Val   Asp Glu Ala Ala Ala   Glu Ala Ala Gly Asp   Ala Tyr Gly Gly
-120   -115   -110

Ser   Val Phe Asp Thr Glu   Ser Leu Glu Leu Thr Val Leu Val Thr Asp
-105   -100   -95   -90

Ala Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu
              -85   -80   -75

Val Ser Tyr Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn

30

```
              -70                      -65                      -60

        Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala
                -55                 -50                 -45


        Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val
                -40                 -35                 -30


        Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val
        -25                 -20                 -15                 -10


        Thr Thr Ser Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu
                        -5          -1  1                   5


        Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn
                10                  15                  20


        Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val
            25                  30                  35


        Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser
        40                  45                  50                  55


        Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe
                        60                  65                  70


        Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr
                    75                  80                  85


        Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser
                90                  95                  100


        Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln
            105                 110                 115


        Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr
        120                 125                 130                 135


        Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr
                        140                 145                 150


        Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly
                    155                 160                 165
```

```
Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
        170                 175                 180


Val Arg Leu Arg Thr
    185



<210>   44
<211>   1164
<212>   DNA
<213>   Artificial sequence

<220>
<223>   Synthetic protease encoding gene


<220>
<221>   CDS
<222>   (1)..(1164)
<223>   Full length protease

<220>
<221>   sig_peptide
<222>   (1)..(81)

<220>
<221>   misc_feature
<222>   (82)..(1164)
<223>   Propeptide

<220>
<221>   mat_peptide
<222>   (577)..(1164)

<400>   44
atg aaa aaa  ccg ctg gga aaa att  gtc gca agc aca gca  ctt ctt        45
Met Lys Lys  Pro Leu Gly Lys Ile  Val Ala Ser Thr Ala  Leu Leu
        -190                 -185                 -180

att tca gtg  gca ttt agc tca tct  att gca tca gca gct  aca gga        90
Ile Ser Val  Ala Phe Ser Ser Ser  Ile Ala Ser Ala Ala  Thr Gly
        -175                 -170                 -165

gca tta ccg  cag tct ccg aca ccg  gaa gca gat gca gtc  tca atg       135
Ala Leu Pro  Gln Ser Pro Thr Pro  Glu Ala Asp Ala Val  Ser Met
        -160                 -155                 -150

caa gaa gca  ctg caa aga gat ctt  gat ctt aca tca gca  gaa gca       180
Gln Glu Ala  Leu Gln Arg Asp Leu  Asp Leu Thr Ser Ala  Glu Ala
        -145                 -140                 -135

gaa gaa ctt  ctt gct gca caa gat  aca gca ttt gaa gtg  gat gaa       225
Glu Glu Leu  Leu Ala Ala Gln Asp  Thr Ala Phe Glu Val  Asp Glu
        -130                 -125                 -120
```

```
gca gcg gca  gaa gca gca gga gat  gca tat ggc ggc tca  gtt ttt        270
Ala Ala Ala  Glu Ala Ala Gly Asp  Ala Tyr Gly Gly Ser  Val Phe
        -115              -110              -105

gat aca gaa  tca ctt gaa ctt aca gtt ctt gtt aca gat gca gca gca      318
Asp Thr Glu  Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala
        -100             -95               -90

gtt gaa gca gtt gaa gca aca gga gca gga aca gta ctt gtt tca tat       366
Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Val Leu Val Ser Tyr
    -85             -80               -75

gga att gat ggc ctt gat gaa att gtt caa gaa ctg aat gca gct gat       414
Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp
-70             -65               -60               -55

gct gtt ccg ggc gtt gtt ggc tgg tat ccg gat gtt gct gga gat aca      462
Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr
        -50               -45               -40

gtt gtc ctt gaa gtt ctt gaa gga tca ggc gca gat gtt tca ggc ctg      510
Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser Gly Leu
        -35               -30               -25

ctg gca gac gca gga gtc gat gca tca gca gtt gaa gtt aca aca tca      558
Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser
        -20               -15               -10

gat caa ccg gaa ctt tat gca gat att att ggc ggc ctg gca tat tat      606
Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala Tyr Tyr
 -5               -1  1               5                 10

atg ggc ggc aga tgc agc gtt ggc ttt gca gca aca aat gca tca ggc      654
Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala Ser Gly
            15              20                25

caa ccg ggc ttt gtt aca gca ggc cat tgc ggc aca gtt ggc aca cca      702
Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Pro
        30              35                40

gtt tca att ggc aat ggc aaa ggc gtt ttt gaa cga agc att ttt ccg      750
Val Ser Ile Gly Asn Gly Lys Gly Val Phe Glu Arg Ser Ile Phe Pro
        45              50                55

ggc aat gat tca gca ttt gtt aga ggc aca tca aat ttt aca ctt aca      798
Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser Asn Phe Thr Leu Thr
    60              65                70

aat ctg gtt tca aga tat aat tca ggc ggc tat gca aca gtt gca ggc      846
Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Ala Thr Val Ala Gly
75              80                85                90

cat aat caa gca ccg att ggc tca gca gtt tgc aga tca ggc tca aca      894
His Asn Gln Ala Pro Ile Gly Ser Ala Val Cys Arg Ser Gly Ser Thr
            95              100               105
```

33

```
aca ggc tgg cat tgc ggc aca att caa gca aga aat caa aca gtt agg     942
Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln Thr Val Arg
            110                 115                 120

tat ccg caa ggc aca gtt tat agt ctg aca aga aca aca gtt tgt gca     990
Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala
            125                 130                 135

gaa ccg ggc gat tca ggc ggc tca tat att agc ggc act caa gca caa    1038
Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln
        140                 145                 150

ggc gtt aca tca ggc ggc tca ggc aat tgc agt gct ggc ggc aca aca    1086
Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr
155                 160                 165                 170

tat tac caa gaa gtt aat ccg atg ctt agt tca tgg ggc ctt aca ctt    1134
Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu
                175                 180                 185

aga aca caa tcg cat gtt caa tcc gct cca                            1164
Arg Thr Gln Ser His Val Gln Ser Ala Pro
            190                 195


<210>   45
<211>   388
<212>   PRT
<213>   Artificial sequence

<220>
<223>   Synthetic Construct

<400>   45

Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu
        -190            -185            -180

Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly
        -175            -170            -165

Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met
        -160            -155            -150

Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala
        -145            -140            -135

Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
        -130            -125            -120

Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe
        -115            -110            -105
```

34

```
Asp Thr Glu  Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala
        -100                -95                -90


Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Val Leu Val Ser Tyr
    -85                 -80                -75


Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp
-70                 -65                -60                     -55


Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr
            -50                 -45                     -40


Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser Gly Leu
        -35                 -30                 -25


Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser
        -20                 -15                 -10


Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala Tyr Tyr
    -5                 -1  1                 5                     10


Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala Ser Gly
            15                  20                  25


Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Pro
            30                  35                  40


Val Ser Ile Gly Asn Gly Lys Gly Val Phe Glu Arg Ser Ile Phe Pro
        45                  50                  55


Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser Asn Phe Thr Leu Thr
    60                  65                  70


Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Ala Thr Val Ala Gly
75                  80                  85                      90


His Asn Gln Ala Pro Ile Gly Ser Ala Val Cys Arg Ser Gly Ser Thr
            95                  100                 105


Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln Thr Val Arg
        110                 115                 120
```

35

```
Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala
    125                 130                 135


Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln
    140                 145                 150


Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr
155                 160                 165                 170


Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu
                175                 180                 185


Arg Thr Gln Ser His Val Gln Ser Ala Pro
                190                 195
```

```
<210>   46
<211>   165
<212>   PRT
<213>   Artificial sequence

<220>
<223>   Shuffled propeptide O-2.19


<220>
<221>   PROPEP
<222>   (1)..(165)

<400>   46
```

```
Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1               5                   10                  15


Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu
                20                  25                  30


Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
            35                  40                  45


Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe Asp
        50                  55                  60


Thr Glu Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65                  70                  75                  80


Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
```

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
            100   105   110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
            115   120   125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
      130   135   140

Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser Asp
145   150   155   160

Gln Pro Glu Leu Tyr
                165

<210>   47
<211>   166
<212>   PRT
<213>   Artificial sequence

<220>
<223>   Shuffled propeptide G-2.73

<220>
<221>   PROPEP
<222>   (1)..(166)

<400>   47

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1   5   10   15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Ser Ser Ala Glu
            20   25   30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
      35   40   45

Ala Ala Ala Gly Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe Asp
   50   55   60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65   70   75   80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
                85   90   95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
                100   105   110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
                115   120   125

Val Val Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
        130   135   140

Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145   150   155   160

Glu Gln Pro Glu Leu Tyr
                165

<210>   48
<211>   166
<212>   PRT
<213>   Artificial sequence

<220>
<223>   Shuffled propeptide G-1.43

<220>
<221>   PROPEP
<222>   (1)..(166)

<400>   48

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1   5   10   15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser Ser Ser Gln
                20   25   30

Ala Glu Glu Leu Leu Asp Ala Gln Ala Glu Ser Phe Glu Ile Asp Glu
        35   40   45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp
        50   55   60

Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
                85   90   95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
            100   105   110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
            115   120   125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
            130   135   140

Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145   150   155   160

Glu Gln Pro Glu Leu Tyr
                165

<210> 49
<211> 166
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide G-2.6

<400> 49

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1   5   10   15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu
                20   25   30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
            35   40   45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp
        50   55   60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ser Ser Ser Val
65   70   75   80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
                85   90   95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
                100   105   110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
                115   120   125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
        130   135   140

Ala Gly Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145   150   155   160

Glu Gln Pro Glu Leu Tyr
                165

<210> 50
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide G-2.5

<220>
<221> PROPEP
<222> (1)..(165)

<400> 50

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1   5   10   15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro Leu Glu
                20   25   30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
        35   40   45

Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe Asp
    50   55   60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65   70   75   80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
              85   90   95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
              100   105   110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
              115   120   125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
              130   135   140

Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Pro Ala Ala
145   150   155   160

Arg Pro Glu Leu Tyr
                    165

<210>   51
<211>   166
<212>   PRT
<213>   Artificial sequence

<220>
<223>   Shuffled propeptide G-2.3

<220>
<221>   PROPEP
<222>   (1)..(166)

<400>   51

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Asp Gly Ala Glu Ala
1   5   10   15

Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro Ala
              20   25   30

Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp
              35   40   45

Glu Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe
              50   55   60

41

```
Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ala Ala
65              70          75              80

Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His
                85              90              95

Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp
            100             105             110

Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr
            115             120             125

Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Tyr Ser Leu
    130             135             140

Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Pro Ala
145             150             155             160

Ala Gln Pro Glu Leu Tyr
                165


<210>  52
<211>  166
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Shuffled propeptide G-1.4


<220>
<221>  PROPEP
<222>  (1)..(166)

<400>  52

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1               5               10              15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser Ser Ser Gln
                20              25              30

Ala Glu Glu Leu Leu Asp Ala Gln Ala Glu Ser Phe Glu Ile Asp Glu
            35              40              45

Ala Ala Ala Ala Ala Ala Ala Asp Ser Tyr Gly Gly Ser Ile Phe Asp
    50              55              60
```

```
Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65                  70                  75                  80


Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
                85                  90                  95


Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
                100                 105                 110


Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
                115                 120                 125


Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
    130                 135                 140


Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145                 150                 155                 160


Glu Gln Pro Glu Leu Tyr
                165


<210>   53
<211>   166
<212>   PRT
<213>   Artificial sequence

<220>
<223>   Shuffled propeptide G-1.2


<220>
<221>   PROPEP
<222>   (1)..(166)


<400>   53

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1               5                   10                  15


Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu
                20                  25                  30


Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
        35                  40                  45
```

```
Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp
    50                  55                  60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ser Ser Ser Val
65                  70                  75                  80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
                85                  90                  95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
            100                 105                 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
        115                 120                 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
    130                 135                 140

Ala Gly Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145                 150                 155                 160

Glu Gln Pro Glu Leu Tyr
            165
```